Proposed Decision Memo for Sleep Testing for Obstructive Sleep Apnea (OSA) (CAG-00405N)

Decision Summary

CMS proposes that the evidence is sufficient to determine that the results of the sleep tests identified below can be used by a beneficiary's treating physician to diagnose OSA and prescribe CPAP therapy, that the use of such sleep testing technologies demonstrates improved health outcomes in Medicare beneficiaries who have OSA and receive the appropriate treatment, and that these tests are thus reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act.

Therefore, we propose that:

1.

    Type I Polysomnography (PSG) is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed attended in a sleep lab facility.

2.

    A Type II or a Type III sleep testing device is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

3.

    A Type IV sleep testing device measuring three or more channels, one of which is airflow, is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

4.

    A sleep testing device measuring three or more channels that include actigraphy, oximetry, and peripheral arterial tone is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

We are soliciting public comments on this proposed decisions pursuant to §1862(l) of the Social Security Act.

Back to Top

Proposed Decision Memo

TO: Administrative File: CAG # Sleep Testing for Obstructive Sleep Apnea (OSA)

FROM: Steve Phurrough, MD, MPA
            Director, Coverage and Analysis Group

            Louis Jacques, MD
            Director, Division of Items and Devices

            Jean Stiller, MA
            Lead Analyst

Ross Brechner, MD, MS (Stat.), MPH
Lead Medical Officer

SUBJECT: Proposed Coverage Decision Memorandum for Sleep Testing for Obstructive Sleep Apnea (OSA) (CAG-)

DATE: December 23, 2008

I. Proposed Decision

CMS proposes that the evidence is sufficient to determine that the results of the sleep tests identified below can be used by a beneficiary's treating physician to diagnose OSA and prescribe CPAP therapy, that the use of such sleep testing technologies demonstrates improved health outcomes in Medicare beneficiaries who have OSA and receive the appropriate treatment, and that these tests are thus reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act.

Therefore, we propose that:

1.

    Type I Polysomnography (PSG) is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed attended in a sleep lab facility.

2.

    A Type II or a Type III sleep testing device is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

3.

    A Type IV sleep testing device measuring three or more channels, one of which is airflow, is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

4.

    A sleep testing device measuring three or more channels that include actigraphy, oximetry, and peripheral arterial tone is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

We are soliciting public comments on this proposed decisions pursuant to §1862(l) of the Social Security Act.

II. Background

We use the abbreviation PSG to refer to polysomnography or a polysomnogram furnished in a sleep laboratory facility. Unless we specifically describe an unattended use, we will always assume in this document that it has been attended. We note that some authors use the abbreviation NPSG to mean nocturnal PSG. We use the abbreviation HST (home sleep test) to refer to unattended multichannel sleep testing or multichannel sleep monitoring typically furnished in the beneficiary's home. However, it does not exclude these tests being performed in other settings to include a sleep lab.

OSA, sometimes referred to as Obstructive Sleep Apnea Hypopnea Syndrome-OSAHS, is associated with significant morbidity and mortality. It is a commonly underdiagnosed condition that occurs in 4% of men and 2% of women (Young et al. 1993). The prevalence increases with age (up to 10% in persons 65 and older), as well as with increased weight. Complications associated with OSA include excessive daytime sleepiness, concentration difficulty, coronary artery disease, and stroke (Kokturk et al. 2005). It is estimated that 10% of patients with congestive heart failure (CHF) have OSA, which is independently associated with systemic arterial hypertension (Caples et al. 2005). Untreated OSA is associated with a ten-fold increased risk of motor vehicle accidents (Teran-Santos et al. 1999). The most common clinical presentation of patients with OSA is obesity accompanied by excessive daytime drowsiness (20% of adults with BMI > 30 have OSA), although other clinical findings associated with OSA include nocturnal choking or gasping, witnessed apneas during sleep, large neck circumference and daytime fatigue.

Of the three different forms of sleep apnea (obstructive, central, or mixed), OSA is the most common. Patients suffering with sleep apnea may literally stop breathing (apnea) for a short period or have decreased breathing (hypopnea), repeatedly during sleep. The apnea episodes often last for a minute or longer, and can occur hundreds of times during a single night's sleep. During the obstructive apnea episodes, either complete or partial obstruction of the airway occurs. The anatomic site of obstruction is thought to be the soft palate, extending to the base of the tongue. When patients with OSA fall asleep, muscles of this region relax to the point of permitting airway collapse and obstruction. When the airway closes, breathing stops and the sleeper awakens to open the airway. Arousals from sleep usually last only a few seconds, but these brief arousals disrupt continuous sleep and prevent persons from reaching deep stages of sleep (e.g., rapid eye movement sleep-REM), which is necessary in order for the body to rest and replenish strength. The patient repeats this cycle throughout the sleep period.

OSA has often been defined by an apnea/hypopnea index (AHI) or respiratory disturbance index (RDI) of $\geq$ 5 events per hour during sleep (when using this less restrictive definition, the prevalence may be as high as 25% of the population) or by a higher threshold e.g. AHI of $\geq$ 15 per hour (the prevalence is approximately 3%). Medicare covers CPAP for the treatment of OSA if the beneficiary has an AHI or RDI $\geq$ 15 events/hour. Medicare also covers CPAP for the treatment of OSA if the beneficiary has a co-morbidity related to OSA and the AHI or RDI is $\geq$ 5 and < 15. The key diagnostic finding in OSA is episodes of airflow cessation or reduction at the nose and mouth despite evidence of continuing respiratory effort.

Other common clinical findings and measurements used by physicians in the diagnosis of OSA include oxygen desaturation, abnormal oxygen desaturation index, arterial pulsatile tone changes, measurement of airflow, measurement of breathing patterns, Multiple Sleep Latency Testing (MSLT), Maintenance of Wakefulness Testing, computerized EEG analysis, autonomic arousal detection, and body movement analysis.

Diagnostic tests for OSA have historically been classified into four types. The most comprehensive is designated Type I attended facility based PSG, which is considered the reference standard for diagnosing OSA. Attended facility based polysomnogram is a comprehensive diagnostic sleep test including at least electroencephalography (EEG), electro-oculography (EOG), electromyography (EMG), heart rate or electrocardiography (ECG), airflow, breathing/respiratory effort, and arterial oxygen saturation (SaO2) furnished in a sleep laboratory facility in which a technologist supervises the recording during sleep time and has the ability to intervene if needed. Overnight PSG is the conventional diagnostic test for OSA. The American Thoracic Society (ATS 1994) and the American Academy of Sleep Medicine (ASDA 1997) have recommended supervised PSG in the sleep laboratory over 2 nights for the diagnosis of OSA and the initiation of CPAP.

Three categories of portable monitors (used both in attended and unattended settings) have been developed for the diagnosis of OSA. Type II monitors have a minimum of 7 channels (e.g., EEG, EOG, EMG, ECG-heart rate, airflow, breathing/respiratory effort, SaO2-this type of device monitors sleep staging, so calculation of apnea/hypopnea index-AHI can be calculated). Type III monitors have a minimum of 4 monitored channels including ventilation or airflow (at least two channels of respiratory movement or respiratory movement and airflow), heart rate or ECG, and oxygen saturation. Type IV devices may measure one, two parameters, three or more parameters but do not meet all the parameters of a higher level device.

Young et al. (1999) note limited capacity to provide PSG testing to all persons with symptoms of OSA due to the high prevalence of OSA. Some studies have noted false-negative rates of 14 to 25% (Le Bon et al. 2000; Littner 2000). And as noted by Klingshott et al. (2000) associates, the measures derived from PSG (e.g., AHI) correlate poorly with major consequences of OSA such as sleepiness and cognitive impairment. Loube et al. (1999) and others have also noted that these measures do not reliably predict the response to the standard therapy for OSA, nasal CPAP.

PSG alternatives have been sought. Predictive algorithms (predictive formulae) to determine optimal CPAP (Flemons et al. 1994; Maislin et al. 1995; Rowley et al. 2000), screening oximetry (Whitlaw et al. 2005; Chiner et al. 1999), attended/unattended home diagnostic apnea monitoring devices (Sériés et al 1993; Golpe et al. 2002; Whitelaw et al. 2005), and questionnaires (e.g., Epworth Sleepiness Scale; Sleep Apnea Clinical Scores) have been developed to help diagnose OSA. Other strategies that have been suggested to reduce the delay, inconvenience and expense associated with sleep studies include split night studies (Yamashiro et al. 1995), partner titration, and home stepwise titration.

A number of treatment approaches have been recommended for patients with OSA, depending on severity of the disorder (e.g., the degree of clinical symptoms), as well as the objective level of nocturnal respiratory and sleep disturbance (e.g., daytime sleepiness or number of obstructive events per hour of sleep). For patients with severe OSA, nasal CPAP is the treatment of choice. Its regular use improves excessive sleepiness, cognitive performance, and quality of life (Jenkinson et al. 1999; Montserrat et al. 2001). In patients with severe OSA who can not tolerate nasal CPAP, surgical procedures (e.g., uvulopalatopharygnoplasty-UPPP, maxillofacial surgery) may be indicated. In patients with mild to moderate OSA, nasal CPAP may be indicated, though conservative measures such as weight reduction, avoidance of alcohol, avoidance of sleeping in a recumbent position, or intra-oral appliances may be better tolerated.
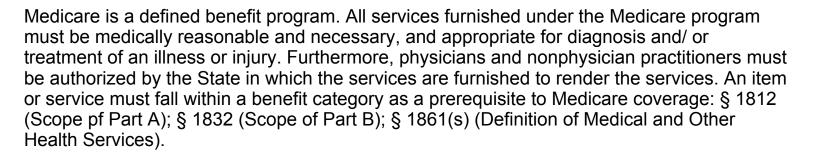
III. History of Medicare Coverage

We received an external request from Itamar Medical requesting a National Coverage Determination (NCD) on whether Home Sleep Testing (HST) devices measuring the peripheral arterial tone (PAT) signal (a measure of sympathetic activation), heart rate, blood oxygen saturation, and sleep time are reasonable and necessary for the diagnosis of obstructive sleep apnea (OSA).

Itamar manufactures the Watch-PAT sleep test device that measures actigraphy, oximetry and peripheral arterial tone. Itamar also asked us to remove this technology from the Type IV classification in the current CPAP NCD and explicitly state that CPAP is covered in beneficiaries diagnosed with OSA using a clinical evaluation and a positive test using this technology.

CMS has addressed the coverage of CPAP is three separate decisions in October, 2001, April 2005, and March 2008. In each of those decisions, we limited coverage of CPAP in patients with OSA to those patients whose diagnosis was based on specific testing modalities. Initially, we limited coverage to OSA diagnosed with PSG. In the latest decision, we expanded coverage to OSA diagnosed with several types of HST. However, we have not, at a national level, specifically addressed coverage of the tests themselves. In other words, we nationally cover CPAP for beneficiaries with OSA if diagnosed with these specific tests; however, coverage of the specific tests is left to local contractor discretion.

Since Watch-PAT is only one of several diagnostic tests for OSA and we do not have an NCD on any of these tests, we have broadened the scope of this NCA to include other sleep test technologies. We have recently reviewed and commented on the evidence available that discusses the benefits of sleep testing for OSA. We are releasing our proposed decision concurrent with opening this NCD.

Benefit Category

Medicare is a defined benefit program. All services furnished under the Medicare program must be medically reasonable and necessary, and appropriate for diagnosis and/ or treatment of an illness or injury. Furthermore, physicians and nonphysician practitioners must be authorized by the State in which the services are furnished to render the services. An item or service must fall within a benefit category as a prerequisite to Medicare coverage: § 1812 (Scope pf Part A); § 1832 (Scope of Part B); § 1861(s) (Definition of Medical and Other Health Services).

Sleep testing to diagnose OSA is considered to be within the following benefit category: §1861(s)(3), diagnostic testing.

The Medicare regulations at 42 CFR 410.32(a) state in part, that "…diagnostic tests must be ordered by the physician who is treating the beneficiary, that is, the physician who furnishes a consultation or treats a beneficiary for a specific medical problem and who uses the results in the management of the beneficiary's specific medical problem."

IV. Timeline of Recent Activities

December 23, 2008	CMS posts a tracking sheet and a proposed decision memorandum on the website and the initial 30 day public comment period begins.

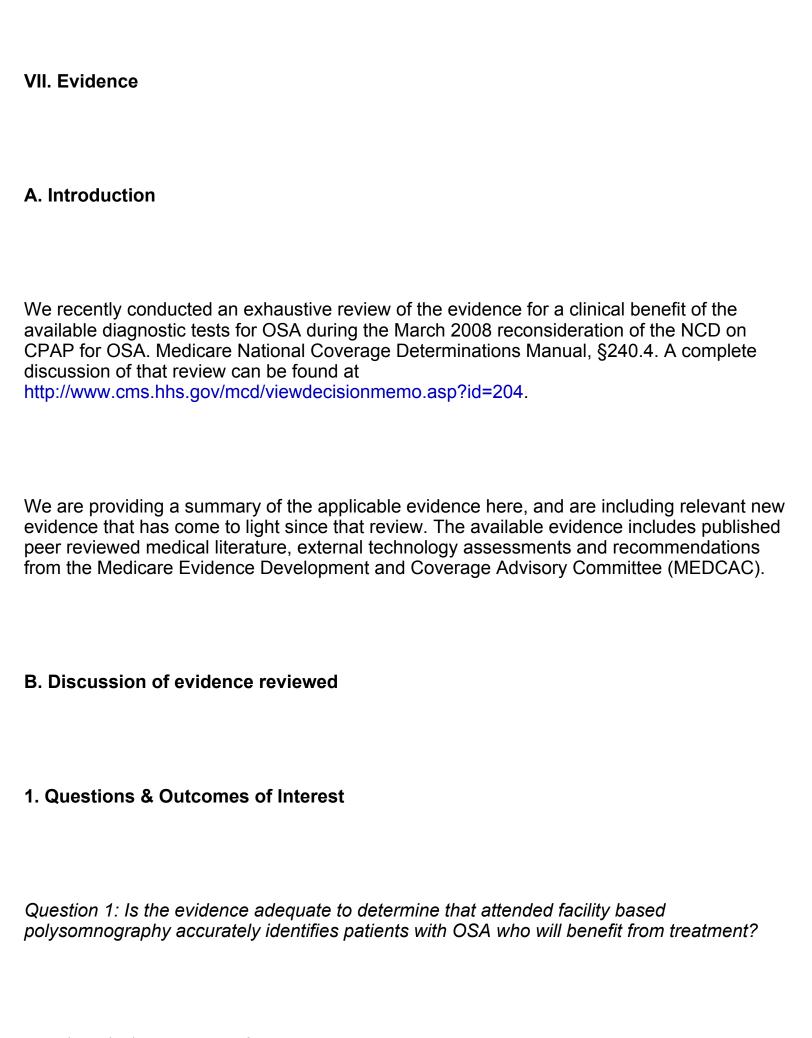## V. Food and Drug Administration (FDA) Status

Certain sleep test devices have been considered and cleared for marketing by the FDA under a 510(k) process.

VI. General Methodological Principles

When making NCDs, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for Medicare beneficiaries. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary under § 1862(a)(1)(A) of the Act.

A detailed account of the methodological principles of study design that are used to assess the relevant literature on a therapeutic or diagnostic item or service for specific conditions can be found in Appendix A. In general, features of clinical studies that improve quality and decrease bias include the selection of a clinically relevant cohort, the consistent use of a single good reference standard, and the blinding of readers of the index test, and reference test results.

Public comment sometimes cites the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

VII. Evidence

A. Introduction

We recently conducted an exhaustive review of the evidence for a clinical benefit of the available diagnostic tests for OSA during the March 2008 reconsideration of the NCD on CPAP for OSA. Medicare National Coverage Determinations Manual, §240.4. A complete discussion of that review can be found at http://www.cms.hhs.gov/mcd/viewdecisionmemo.asp?id=204.

We are providing a summary of the applicable evidence here, and are including relevant new evidence that has come to light since that review. The available evidence includes published peer reviewed medical literature, external technology assessments and recommendations from the Medicare Evidence Development and Coverage Advisory Committee (MEDCAC).

B. Discussion of evidence reviewed

1. Questions & Outcomes of Interest

Question 1: Is the evidence adequate to determine that attended facility based polysomnography accurately identifies patients with OSA who will benefit from treatment?
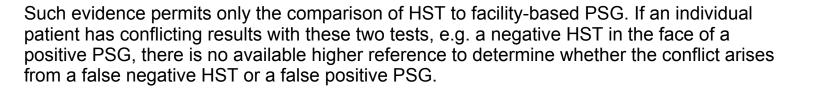
Question 2: For which unattended out of facility sleep test technologies is the evidence adequate to determine that sleep testing accurately identifies patients with OSA who will benefit from treatment?

As diagnostic tests, PSG and HST would not be expected to directly change health outcomes. Rather, a diagnostic test affects health outcomes through changes in disease management brought about by physician actions taken in response to test results. Such actions may include decisions to treat or withhold treatment, to choose one treatment modality over another, or to choose a different dose or duration of the same treatment. To some extent the usefulness of a test result is constrained by the available treatment options. As noted in the Background section, the number of practical treatment options for OSA is limited. Most patients get CPAP; a few get oral appliances or surgery. A patient whose OSA is not readily controlled with CPAP may seek other treatment, continue CPAP with lesser benefit, or discontinue CPAP and not seek further medical treatment. In addressing the questions above, one of the factors we considered is whether there is sufficient evidence that the incremental information derived from PSG or HST leads to improved treatment of OSA by causing physicians to prescribe a different treatment than they would have prescribed without access to the test results.

Outcomes of interest for a diagnostic test are not limited to determining its accuracy but also include beneficial or adverse clinical effects, such as changes in management due to test findings or preferably, improved health outcomes for Medicare beneficiaries. Ideally, we would see evidence that the systematic incorporation of PSG or HST results into a treatment algorithm leads treating physicians to prescribe different and better treatment than they would otherwise have prescribed, and that those patients whose treatment is changed by test results remain on the regimen and achieve better long term OSA control documented by repeated assessments over time.

There is no anatomic or physiologic "gold standard" for the diagnosis of obstructive sleep apnea, in contrast to conditions such as cancer where a tissue biopsy result is the definitive standard reference. In studies that compare HST to facility-based PSG, the investigators have used the PSG result as the standard reference; i.e. the PSG result is used to define the true disease state for the individual patient. This is less than ideal since the true sensitivity and specificity of PSG in diagnosing OSA is not well documented and this deficiency poses a practical difficulty in diagnosing OSA. Given the absence of a true "gold standard" reference, the clinical application of terms such as sensitivity and specificity is not straightforward.

Such evidence permits only the comparison of HST to facility-based PSG. If an individual patient has conflicting results with these two tests, e.g. a negative HST in the face of a positive PSG, there is no available higher reference to determine whether the conflict arises from a false negative HST or a false positive PSG.

2. External technology assessments

Systematic reviews are based on a comprehensive search of published studies to answer a clearly defined and specific set of clinical questions. A well-defined strategy or protocol (established before the results of the individual studies are known) guides this literature search. Thus, the process of identifying studies for potential inclusion and the sources for finding such articles is explicitly documented at the start of the review. Finally, systematic reviews provide a detailed assessment of the studies included. CMS commissioned two TAs from AHRQ for the March 2008 CPAP NCD reconsideration:

-

- Home diagnosis of Obstructive Sleep Apnea-Hypopnea Syndrome, and
- Obstructive Sleep Apnea-Hypopnea Syndrome: modeling different diagnostic strategies

We summarize them below. The full reports are available at the following CMS website: http://www.cms.hhs.gov/mcd/viewtechassess.asp?id=204.

Home diagnosis of Obstructive Sleep Apnea-Hypopnea Syndrome

Ninety-three studies were included in a review of the literature. Eligible studies assessed the ability of sleep studies at baseline to predict response to CPAP treatment or CPAP use, the comparison of measurements with portable monitors and facility-based PSG, and the safety of sleep studies.

The TA reported that the reference standard for the diagnosis of OSAHS is facility-based PSG, a comprehensive sleep study that records and evaluates a variety of cardiorespiratory and neurophysiologic signals during sleep time. It quantifies the severity of disturbances with the Apnea-Hypopnea Index (AHI). Higher AHI values imply more severe sleep disturbances. Typically, a value of 15 or more events/hour of sleep is considered to be suggestive of OSAHS. An AHI suggestive of OSAHS is neither sufficient nor necessary for the diagnosis of the condition, as the severity of symptoms has to be accounted for, and other conditions affecting sleep may need to be excluded. Baseline AHI is only modestly associated with response to CPAP use among people with high (pre-test) probability for OSAHS. The same is true for other indices obtained from sleep studies such as the mean or minimum O_2 saturation, apnea index, hypopnea index, frequency of arousals and other quantities.

Based on limited data, the authors conclude that type II monitors may identify AHI suggestive of OSAHS with high positive likelihood ratios (> 10) and low negative likelihood ratios (< 0.1) both when the portable monitors were studied in the sleep laboratory and at home. Type III monitors may have the ability to predict AHI suggestive of OSAHS with high positive likelihood ratios and low negative likelihood ratios for various AHI cutoffs in laboratory-based PSG, especially when manual scoring is used. The ability of type III monitors to predict AHI suggestive of OSAHS appears to be better in studies conducted in the specialized sleep unit compared to studies in the home setting. Some studies of type IV monitors also showed high positive likelihood ratios and low negative likelihood ratios, at least for selected sensitivity and specificity pairs from ROC curve analyses. As with type III monitors, the ability of type IV monitors to predict AHI suggestive of OSAHS appears to be better in studies conducted in specialized sleep units. Medicare beneficiaries are older than the studied subjects (the median average age was approximately 50 years in the analyzed studies), and may more often have conditions other than OSAHS that affect sleep (e.g., Periodic Limb Movements in Sleep and Restless Leg Syndrome; cardiac insufficiency). These conditions may be misdiagnosed as OSAHS by sleep monitors that do not record channels necessary for the differential diagnosis of OSAHS. Therefore, some type III and type IV monitors may yield more false positives among Medicare beneficiaries, compared to what was observed in the assessed studies. For studies in the home setting, there are no direct data on whether and to what extent technologist support and patient education affect the comparison of portable monitors with facility-based PSG.

For monitors that may be considered other than Type II, III, or IV, the authors found there is insufficient evidence to judge their value in diagnosing OSA. The TA differentiated Type IV monitors with three or more channels from those with one or two channels, finding greater diagnostic ability for the former. We note that the TA reviewed the Watch-PAT100 device as a Type IV device with three or more channels.

Obstructive Sleep Apnea-Hypopnea Syndrome: modeling different diagnostic strategies

The TA authors created a model to test the impact of different OSA strategies. When middle-aged people (50 years old) with symptoms and signs suggestive of OSAHS are tested in the home setting, approximately 10 percent of those with OSAHS are expected to remain undiagnosed; approximately 15 percent of those without OSAHS receive false-positive diagnoses. For older adults (70 years old), the expected number of misclassifications is larger, due to the expected increase in false positive diagnoses (30 percent). With the combination strategy that uses home diagnosis and split-night PSG, almost 20 percent of middle-aged people with OSAHS received a (false) negative diagnosis, while the proportion of false positive results among 50 year-old people without OSAHS was very low (1 percent). The expected numbers were similar among older adults (70 years old).

Both for middle-aged people and for older adults, the average time spent undiagnosed is practically negligible for the strategies that use home monitoring. In the combination strategy, people with positive diagnosis with the portable monitors receive a final split-night PSG diagnosis within 15 weeks on average.

When diagnosis of OSAHS and treatment initiation are managed outside the sleep laboratories in the home setting, middle-aged people with OSAHS spend on average 10 weeks or 9 percent of the total follow-up time in undiagnosed health states. Significantly, the corresponding mean time delay for middle-aged people is 27 weeks when they are managed with facility-based PSG. This number mainly reflects those with false negative diagnoses, who are never started on CPAP. The same delay is expected among older adults (70 years old).

With the combination strategy, using home diagnosis and split-night PSG, correctly diagnosed people initiate CPAP after approximately 15 weeks. However, one fifth of the patients are not diagnosed and, overall, the average time spent while not on CPAP ("high-risk" states) becomes 33 weeks. Similar numbers are expected among older adults who have OSAHS.

3. Internal technology assessment

<u>Literature Search</u>
CMS performed an extensive literature search utilizing PubMed for randomized controlled trials (RCTs), systematic reviews, and series studies evaluating the technology used for the diagnosis of OSA. The literature search was limited to humans

There are currently several proposed mechanisms to diagnose OSA and determine the need for and benefit of OSA treatment, specifically CPAP. These include clinical diagnosis alone, PSG, home testing with various devices and a diagnosis made by using a trial of CPAP.

<u>Clinical Diagnosis Alone and Clinical Diagnosis with PSG</u>

<u>Crocker et al. (1990)</u> studied whether the number of PSGs required for diagnosis of OSA could be reduced in the population. They enrolled 100 consecutive patients (average age 50) screened by family and sleep physicians. The patients were then tested by PSG. A clinical model was created for predicting a diagnosis of OSA as compared to PSG and was applied to the next 114 consecutive patients. The model correctly classified 33 of 36 persons with OSA by correctly predicting an AHI > 15 and it correctly classified 35 of 69 patients by correctly predicting an AHI $\leq$ 15. In the model, BMI, reported apnea, age, and hypertension were statistically significant factors. The model had a sensitivity of 92% for predicting OSA when compared to PSG and a specificity of 51%. The authors concluded that clinical observation might reduce the need for PSG in the diagnosis of OSA by one-third.

Deegan et al. (1996) compared the predictive value of certain clinical features to PSG for a diagnosis of OSA. Two hundred fifty consecutive patients (average age 45) were pre-screened by a physician and had a clinical assessment and administration of a sleep questionnaire, along with PSG. One hundred thirty six (54%) had an AHI $\geq$ 15 (considered positive for a diagnosis of OSA) and 114 (46%) had an AHI < 15 (not considered positive for OSA). Using clinical features and oximetry, 32.4% of patients could be confidently categorized, compared to PSG, as either having a true diagnosis of OSA or not having OSA. Significant factors in the model were BMI, alcohol intake, and age. The authors concluded that clinical observation may reduce the need for PSG in the diagnosis of OSA by approximately one-third.

Haponik et al. (1984) asked whether or not PSG is necessary to assess the presence and severity of sleep-disordered breathing. They enrolled 37 patients (average age 50) with clinically suspected OSA, administered a questionnaire and did PSG testing. Compared to PSG (AHI $\geq$ 15 as cutoff for positive diagnosis of OSA) the clinical testing information had a sensitivity of 64% for a correct diagnosis of OSA and a specificity of 100%. The authors concluded that a single, brief clinical observation alone is an ineffective screening procedure for detecting OSA.

Julià-Serdà et al. (1984) enrolled 225 consecutive referrals to a sleep clinic (average age 45 in the non-OSA group and 52 in the OSA group) with suspected OSA to determine whether or not cephalometry was useful in sparing PSG. All subjects had clinical assessment with an ESS questionnaire, physical exam and history. In addition they also had spirometry, cephalometry, and PSG testing. A statistical model was built to estimate a patient's probability of a correct diagnosis of OSA as compared to PSG (using a cutoff value of AHI $\geq$ 10), based on clinical variables, physical examination, pulse oximetry, cephalometry, and soft palate and uvula measurements. The sensitivity of the model for a correct diagnosis of OSA as compared to PSG was 93% and the specificity was 83%. The authors concluded that cephalometry plus oximetry plus history and physical exam is capable of sparing the need for PSG in diagnosing OSA.

Dixon et al. (1997) attempted to create a clinical model for predicting a correct diagnosis of OSA as compared to PSG in 99 pre-operative Laparoscopic Adjustable Gastric Banding patients with average age in their four groups ranging from 35 to 44. A thorough sleep history and physical examination were performed, checking for symptoms such as nocturnal choking, waking unrefreshed, morning headaches, excessive daytime sleepiness and poor sleep quality. An ESS was administered and all patients had a PSG test. The PSG was hand scored. For a PSG cutoff of AHI $\geq$ 15, independent predictors for a diagnosis of OSA were observed sleep apnea (the only positive symptom predictor of an AHI $\geq$ 15), male sex, higher BMI, age, fasting insulin and glycosylated hemoglobin A1c. From the model created, a scoring mechanism was established and a score of > 3 had a sensitivity of 89% for a correct diagnosis of OSA as compared to PSG and a specificity of 81% for moderate/severe OSA. The authors concluded that a simple method of predicting OSA in severely obese symptomatic subjects can assist in limiting the use of PSG to those with greater risk

Lim et al. (2006) performed a study to determine if a clinical model could be developed to predict OSA diagnosis from clinical diagnosis only. Seventy-one consecutive snorers (average age 44) referred for an evaluation for OSA were enrolled. OSA status was determined by clinical assessment based on symptoms suggestive of OSA as well as an ESS and BMI measurement. A PSG was administered and a clinical assessment model was created and used in identifying the 'non-apneic snorers' among patients referred with snoring. The model made use of the ESS score (using a cutoff of $\geq$ 15), the BMI (using a cutoff of $\geq$ 28), and the presence of symptoms such as nocturnal choking, witnessed apnea, daytime hypersomnolence or morning headaches. Compared to PSG using a cutoff of AHI > 10, the model had a sensitivity of 93.4% and a specificity of 60% for correctly diagnosing OSA. The authors concluded that identifying 'non-apneic snorers' in whom PSG could be avoided can be correctly accomplished via a clinical assessment if two out of three of the following are absent: 1) ESS score $\geq$ 15; 2) a BMI $\geq$ 28; and 3) the presence of specified symptoms such as nocturnal choking, witnessed apnea, daytime hypersomnolence or morning headaches.

Hoffstein et al. (2006) utilized data from 594 patients with an average age 47 who were referred to sleep clinic for suspicion of sleep apnea and were all seen by the same physician to determine if it was possible to develop a clinical model to predict a correct diagnosis of OSA from a clinical exam. A PSG with a cutoff of AHI > 10 was used for the diagnosis of OSA. The independent predictors of a correct diagnosis of OSA as compared to PSG were age, sex, BMI, partner observation of apnea and pharyngeal exam findings (normal vs abnormal). Compared to PSG, the subjective (clinical) impression alone showed a sensitivity of 63% for a correct diagnosis of OSA and a specificity of 60%. The authors concluded that subjective impression alone is not enough to reliably identify patients with or without a correct diagnosis of OSA as compared to PSG.

Garcia et al. (2003) studied whether or not they could predict a correct diagnosis of OSA with a clinical model. They enrolled 227 consecutive patients (average age 58) measuring clinical signs and symptoms and performing a PSG. They then took the next 102 patients and tested their model for clinical diagnosis of OSA (total 329). They utilized an AHI > 30 as a cutoff for a correct diagnosis of OSA. In the model created, they utilized a cut point of 11 for the ESS and of 30 for BMI and included other significant and independent factors of age, sex, BMI, neck circumference history and the referring physician's subjective feeling (dichotomized into 'yes' or 'no') as to each patient's probability of having an AHI > 30. Compared to PSG, the model had a sensitivity of 80% for a correct diagnosis of OSA and a specificity of 93%. The authors concluded that prior to diagnostic tests for OSA; clinical data can be useful for identifying patients suspected of having AHI > 30.

Kushida et al. (1997) attempted to predict OSA with a morphometric predictor model. Thirty patients (age range 15-75) were used to create the model and the model was then prospectively tested on the first consecutive 300 of a total of 423 patients referred for a diagnosis of OSA. All patients were also tested with PSG using a cutoff of AHI > 10. The regression model included oral cavity measurements of the palatal height by two separate calipers measuring the distance between the mesial surfaces of the crowns of the second molars to obtain either the maxillary intermolar distance or the mandibular intermolar distance. BMI and neck circumference measurements were also made. The morphometric model had a sensitivity of 97.6% for a correct diagnosis of OSA as compared to PSG and a specificity of 100%. The authors concluded that the model may be clinically useful as a screening tool for OSA rather than as a replacement for PSG.

Pillar et al. (1994) compared a clinical diagnosis of OSA to PSG (cutoff AHI > 10). Eighty-six patients (average age 47) referred to a sleep clinic for suspicion of OSA were enrolled. The authors did not mention whether or not the subjects were consecutively enrolled. All patients answered a detailed sleep questionnaire, had a brief physical examination and had PSG testing. Compared to PSG (cutoff AHI > 10), a clinical diagnosis of OSA had a sensitivity of 79% and a specificity of 50%. With regards to the model, the independent factors for a true diagnosis of OSA were neck circumference, age, self reporting of apnea and falling asleep unintentionally. Compared to PSG, the sensitivity was 92% and the specificity was only 18%. The authors concluded that clinical evaluation cannot replace PSG.

Rauscher et al. (1993) enrolled 98 habitual snorers and 89 patients (average age 58 overall) with a positive diagnosis of OSA by PSG to see which snorers referred to a sleep laboratory need PSG for the diagnosis of OSA. A regression model was created that included weight, height, sex, witnessed episodes of apnea and falling asleep reading. This model was applied to 116 consecutive patients referred for investigation of heavy snoring. All patients with negative oximetry and a probability value < 0.31 for having OSA had an AHI < 10 by PSG. The authors concluded that snorers with negative oximetry classified as not having OSA by this model do not need PSG.

Viner et al. (1994) examined whether or not history and physical examination can predict a correct diagnosis of OSA as compared to PSG. They enrolled 410 patients (average age 50) referred for clinically suspected OSA. They conducted a blinded comparison of history and physical examination versus results of nocturnal PSG utilizing a cutoff point of AHI $\geq$ 10. The regression model created included as significant independent factors age, BMI, sex, witnessed episodes of apnea and falling asleep reading. They noted that for p < 0.20 (a predicted probability of less than 20% of having OSA) the clinical model had 94% sensitivity and 28% specificity of correctly predicting a diagnosis of OSA as compared to PSG. Subjective impression alone had a sensitivity of 52% and a specificity of 70% for correctly predicting a diagnosis of OSA as compared to PSG. The authors concluded that in patients with a low predicted probability of having a correct diagnosis of OSA, approximately one-third do not need a PSG for diagnosis.

Home testing for OSA
Types II, III & IV
CMS reviewed the AHRQ TA assessment above and also found some additional evidence on HST.

Tsai et al. (2002) performed a study to create a decision rule for diagnostic testing in OSA. They enrolled 75 patients (average age 47) referred to a sleep clinic for suspicion of sleep apnea. No mention of consecutive selection was made. Each patient had portable RDI testing (using a cutoff of RDI $\geq$ 10) and nocturnal oxygen saturation measurements. During the feasibility phase, patients underwent routine clinical assessment plus the upper airway physical examination protocol (UAPP), performed by two investigators. Unreliable or time consuming measurements were eliminated from the UAPP based on clinical judgments and history of snoring and body position based on the consensus of the two investigators. A decision rule was developed using three predictors: a cricomental space (the perpendicular distance between the midpoint of the cricomental line, a straight line from the chin to the cricothyroid cartilage, and the skin of the neck) of 1.5 cm or less, a pharyngeal grade (I = palatopharyngeal arch intersects at the edge of the tongue; II = palatopharyngeal arch intersects at 25% or more of the tongue diameter; III = palatopharyngeal arch intersects at 50% or more of the tongue diameter; IV = palatopharyngeal arch intersects at 75% or more of the tongue diameter) of more than II and the presence of overbite. For patients with all 3 predictors (17%), the decision rule had a PPV of 95% and an NPV of 49% for a true diagnosis of OSA by PSG. Comparable performance was obtained in a validation sample of 50 patients referred for diagnostic testing. The authors concluded that their decision rule provides a simple, reliable and accurate method of identifying a subset patients with and, perhaps more importantly, without a true diagnosis of OSA.

Ayappa et al. (2008) evaluated the ARES Unicorder, a self-applied, limited-channel portable monitoring device for the evaluation of sleep disordered breathing (SDB) using a prospective study with blinded analysis. Eighty patients with suspected OSA and 22 volunteers were enrolled. Interventions used the ARES™ Unicorder at home for 2 nights using only written instructions. The number of men in the suspected OSA group was 60 and the number of women was 17, while in the volunteer group it was 9 and 11 respectively. The mean age in the suspected OSA group was 46 (range 26-74), while the mean age in the volunteer group was 36 (range 19-73). The mean BMI in the suspected group was 30 (range 21-70) the mean BMI was 24 in the volunteer group (range 19-32).

Within 2 weeks, they returned to the laboratory for full nocturnal polysomnography (NPSG) with simultaneous monitoring with the Unicorder. NPSGs were scored manually to obtain an apnea-hypopnea index based on Medicare guidelines (AHI4%) and a respiratory disturbance index (RDI). ARES studies were autoscored and reviewed to obtain indices based on equivalent definitions i.e., AHI4%ARES, and apnea hypopnea (events with 1% desaturation) index (AHI1%ARES). Indices from the NPSG were compared to the in-lab ARES and in-home ARES indices using mean differences and the intraclass correlations (ICC).

For the in-lab comparison, there was high concordance between AHI4%NPSG and AHI4%ARES (ICC = 0.96, mean difference = 0.5/hour) and RDINPSG and AHI1%ARES (ICC = 0.93, mean difference = 3.2/hour). For NPSG versus In-Home ARES comparison, there was good concordance between AHI4%NPSG and AHI4%ARES (ICC = 0.8, mean difference = 4.1/ hour) and RDINPSG and AHI1%ARES (ICC = 0.8 mean difference = 8.6/ hour). The diagnostic sensitivity of in-lab ARES™ for diagnosing OSA using an RDI cut-off of 15 per hour was 95% and specificity was 94%, with a positive likelihood ratio (LR+) = 17.04, and negative likelihood ratio (LR-) = 0.06. For in-home ARES data the sensitivity was 85% and specificity 91% (LR+ = 9.34, LR- = 0.17). There was good agreement between the manually scored NPSG OSA indices and the autoscoring ARES algorithm.

The authors concluded that the ARES Unicorder provides acceptably accurate estimates of OSA indices compared to conventional laboratory NPSG for both the simultaneous and in-home ARES data and that the high sensitivity, specificity, and positive and negative likelihood ratios obtained in the group they studied supported the utility of an ambulatory limited-monitoring approach not only for diagnosing sleep disordered breathing but also to rule out OSA in suitably selected groups.

Alvarez et al. (2008) aimed at evaluating the reliability of home respiratory polygraphy (using the Edentec Monitoring System polygraph) for the diagnosis of sleep apnea–hypopnea syndrome (OSA) and comparing the cost of this technique with that of nighttime polysomnography performed in a sleep laboratory. Using a prospective study design with a random sample of patients with clinically suspected OSA, the participants underwent both home respiratory polygraphy and nighttime PSG and were blinded as to the results of their first test. The study population was composed of 45 patients with a mean (SD) age of 52.3 (11) years of whom 21 (46.6%) were diagnosed with OSA, defined by an AHI> 10 by nighttime PSG. Comparison of the results between PSG and home polygraphy revealed statistically significant correlations for all comparisons. The optimal cutoff in this population was a RDI of 13.7 or more, for which the area under the receiver operating characteristic curve was 87.5% (95% confidence interval, 74.2%-95.4%). The authors concluded that home respiratory polygraphy is a reliable technique for the diagnosis of OSA and that uncertain results must be verified by nighttime PSG.

Oximetry

Both PSG and HST have an oximetry component, which monitors oxygen desaturation. A number of authors have claimed that just using the oximetry component alone can help in making a diagnosis of OSA (Nuber et al. 2000; Sériés et al 2005; Sériés et al.1993; Guylay et al. 1993).

Guylay et al. (2006) studied 98 non-consecutive patients referred for suspicion of sleep apnea to a sleep clinic to compare clinical assessment with home oximetry in the diagnosis of OSA. All patients answered a questionnaire, had a history and physical exam, and had PSG testing using a cutoff value of AHI $\geq$ 15 for diagnosis of OSA. Physicians also independently estimated the likelihood of their patient having a true diagnosis of OSA on PSG testing. Compared to PSG, the independent clinical (physician) assessment had a sensitivity of 79% and a specificity of 50% for correctly diagnosing OSA at the cutoff value of AHI $\geq$ 15. Compared to PSG, oximetry with a desaturation of 2% had a sensitivity of 65% and a specificity of 74% for diagnosing OSA at the cutoff value of AHI $\geq$ 15. For desaturations of 3%, the corresponding sensitivity and specificity were 51% and 90%, respectively. If the percentage of sleep time spent at $SaO2 < 90$ was $\geq$ 1%, the sensitivity for a true diagnosis of OSA as compared to PSG (AHI $\geq$ 15) was 93% and the specificity was 51%. The authors concluded that being at $SaO2 < 90$ for < 1% of the time on home oximetry practically excludes OSA.

As noted above, a number of studies have shown that oximetry measurement helps the diagnostic accuracy of OSA. Sériés et al. (1993) performed one of the earliest studies exploring this relationship. Using 240 consecutive patients with a confirmed (AHI > 10 on PSG) diagnosis of OSA (all were clinically suspected of having OSA because of loud snoring; nocturnal choking and awakenings or apneic events or all three reported by a bedmate; bad sleep quality; and daytime hypersomnolence), they found that oximetry had a 98% sensitivity for diagnosing OSA (AHI > 10), but a specificity of only 48%.

Magalang et al. (2003) explored the relationship between oximetry and OSA. They noted that several quantitative indices derived from overnight pulse oximetry have been used to predict the presence of OSA: (1) number of episodes of oxyhemoglobin desaturations below a threshold-usually a 3% or 4% decline below baseline, (2) the cumulative time spent below an oxyhemoglobin saturation of 90%, and (3) the Δ [delta] index—a measure of the variability of the oxyhemoglobin saturation. The researchers wanted to compare these indices and determine if some combination of these indices predicted an individual's AHI as measured by PSG. Using a derivation group which consisted of 224 consecutive patients, a prediction model was generated based on AHIs from the calculated quantitative indexes. The model was further validated using two groups of consecutive eligible patients (group 1 consisted of 101 patients and group 2 consisted of 191 patients). All patients underwent standard overnight PSG and measurement of arterial oxyhemoglobin (by pulse oximeter).

The major findings of the study revealed that among the different oximetry indices, the Δ index was the best predictor of the presence of OSA, though the number of desaturation events provided similar levels of diagnostic accuracy (sensitivity of a Δ index of > 0.63 in the diagnosis of OSA was 91%, while the specificity was 59%). An aggregation of the model using combinations of all oximetry indices reduced the prediction error (r^2 = 0.70, $p < 0.05$) compared to using the Δ index alone (r^2 = 0.60), improving the precision of prediction of the AHI. The correlation between the predicted and actual AHI was 0.77 when using the Δ index alone, but improved to 0.83 when using a combination of all three oximetry indices. The authors note that one limitation of the study is that the prediction model was validated using overnight pulse oximetry obtained simultaneously with PSG data in the sleep laboratory. However, one advantage of this approach is it eliminated the potential confounder of night-to-night variability of AHI, as well ensuring that oximetry data were collected in exactly the same environment as the PSG data.

Vazquez et al. (2000) studied the diagnostic performance of an automated digital oximetry analysis based on falls and recovery of oxygen saturation and compared the results to PSG. After excluding subjects not eligible for the study, 241 participants with suspected OSA were enrolled in the study and randomly assigned to either PSG or automated off-line analysis of the digitally recorded oximetry signal. Study outcomes included PSG-derived AHI, and oximeter-derived respiratory disturbance index (RDI). The study revealed that the PSG-derived AHI and the oximetry-derived RDI were strongly correlated (R = 0.97); the mean (± 2SD) of the differences between AHI and RDI was 2.18 (± 12.34)/h. Using a case definition of 15 episodes/hour for both AHI and RDI, the sensitivity and specificity were 98% and 88% respectively. The authors noted that one limitation of the applicability of this study was that the algorithm was evaluated by comparison with simultaneous PSGs. They also commented that a number of studies have shown a difference in RDI between home and hospital settings, despite using the same monitor and controlling for technical difficulty. But the authors were quick to note that by evaluating patients in the sleep laboratory, potential confounders (such as technical difficulties associated with remote monitoring, night-to-night variability, and the effects of the home environment on RDI) are eliminated.

Devices measuring peripheral arterial tone, actigraphy and oximetry

We were asked to perform a separate review of the Watch-PAT100 device, as there has been some uncertainty expressed about how to classify this device in the current Type schema. Watch-PAT100 is an HST device which measures the peripheral arterial tone (PAT) and actigraphy (a measure of movement) which are recorded with an ambulatory wrist-worn device (Watch-PAT100). The PAT signal is a measure of the pulsatile volume changes at the finger tip reflecting sympathetic tone variations. The algorithm was developed using a training set of 30 patients recorded simultaneously with polysomnography and Watch-PAT100. The WATCH-PAT100 indirectly detects apnea/hypopnea events by identifying surges of sympathetic activation associated with the termination of these events. This information is further combined with heart rate and pulse oximetry data that are analyzed by the automatic algorithm of the system. This detects respiratory events and calculates the PAT RDI (PRDI).

We found 20 separate articles, papers, editorials, and fact sheets addressing this technology. Of these, CMS determined that 13 were not relevant due to qualities pertaining to sample size, type of evidence, having not been published in a peer reviewed journal or not relevant to this data needed for this NCD. The remaining 9 are reviewed below.

Pittman et al. (2004) aimed at assessing the accuracy of a wrist-worn device (Watch-PAT 100) to diagnose obstructive sleep apnea in the home. Participants were not consecutive patients but were a sample of patients who disclosed on a comprehensive questionnaire between June and December of 2002 that they were interested in being contacted about research studies conducted at the sleep laboratory. All thirty subjects completed 2 overnight diagnostic studies with the test device: 1 night in the laboratory with concurrent polysomnography and 1 night in the home with only the Watch-PAT100. The mean age of these subjects was 43.2 ± 10.8 years and mean body mass index was 33.9 ± 7.1 kg/m2. The mean Epworth Sleepiness Scale score was 9.2 ± 4.7 (range 2-18). The order of the laboratory and home study nights was random.

The frequency of respiratory events on the PSG was quantified using indexes based on 2 definitions of hypopnea: the respiratory disturbance index (RDI) using American Academy of Sleep Medicine (AASM) Task Force criteria for clinical research, and the Medicare guidelines. The PRDI and oxygen desaturation index (PAT ODI) were then evaluated against the polysomnography AASM guidelines (RDI.C) and Medicare guidelines (RDI.M), respectively, for both Watch-PAT100 diagnostic nights, yielding in-lab and home comparisons. The setting for the PSGs was a sleep laboratory affiliated with a tertiary-care academic medical center. The PDG and PAT measures were compared using the mean [2 SD] of the differences and the intra-class correlation coefficient (ICC). The receiver-operator characteristic curve was used to assess optimum sensitivity and specificity and calculate likelihood ratios. For the in-lab comparison, there was high concordance between: RDI.C and PAT RDI: ICC = 0.88, mean difference 2.5 [18.9] events per hour RDI.M and PAT ODI: ICC = 0.95, mean difference 1.4 [12.9] events per hour sleep time: ICC = 0.70, mean difference 7.0 [93.1] minutes. For the home-laboratory comparison, there was good concordance between: RDI.C and PAT RDI: ICC = 0.72, mean difference 1.4 [30.1] events per hour RDI.M and PAT ODI: ICC = 0.80, mean difference 1.6 [26.4] events per hour. Home studies were performed with no technical failures.

The authors concluded in this study of a population of 30 patients suspected of having obstructive sleep apnea that the Watch-PAT100 can quantify an ODI that compares very well with Medicare criteria for defining respiratory events and an RDI that compares favorably with AASM criteria for defining respiratory events. They further believe that the device can be used with a low failure rate for single use in the lab and home for self-administered testing.

Zou et al. (2006) aimed at assessing the accuracy of a portable monitoring device based on PAT to diagnose obstructive sleep apnea (OSA) and to propose a new standard for limited-channel device validation using synchronized polysomnography (PSG) home recordings in a population-based cohort, i.e. in a population sample not preselected for OSA symptoms. The 98 subjects (55 men; age, 60 ± 7 year; body mass index, 28 ± 4 kg/m^2) from a community of 18,000 in Sweden had single-night, unattended PSG and Watch-PAT100 in the home. They were consecutively recruited from the Swedish Skaraborg Hypertension and Diabetes Project. The accuracy of the algorithms used for AHI and RDI calculation from Watch-PAT100 testing were mainly based on 2 components: the oxygen-saturation data plus an indication of autonomic activation from the PAT signal. Events for AHI and RDI calculation were defined as follows: (1) any oxygen desaturation event of 3% or more was counted into both the AHI and RDI and (2) a respiratory event detected from the PAT signal was based on a PAT-signal attenuation that was coupled with pulse-rate acceleration. Watch-PAT100 measurements on RDI, AHI, ODI, and sleep-wake detection were cross walked and compared with PSG data taken from simultaneous PSG recordings.

The mean PSG-AHI in this population was 25.5 ± 22.9 events per hour. The Watch-PAT100 RDI, AHI, and ODI correlated closely (0.88, 0.90, and 0.92; p < .0001, respectively) with the corresponding indexes obtained by PSG. The areas under the curve for the receiver-operator characteristic curves for Watch-PAT100 AHI and RDI were 0.93 and 0.90 for the PSG-AHI and RDI thresholds of 10 and 20 (p < .0001) respectively. The agreement of the sleep-wake assessment was 82 ± 7%. The authors concluded that the Watch-PAT100 was reasonably accurate for unattended home diagnosis of OSA in a population sample not preselected for OSA symptoms. The authors propose that simultaneous home PSG recordings in population-based cohorts is a reasonable validation standard for assessment of simplified recording tools for OSA diagnosis.

Pillar et al. (2002) state that arousals from sleep are associated with increased sympathetic activation and are therefore associated with peripheral vasoconstriction. The authors hypothesized that digital vasoconstrictions as measured by peripheral arterial tonometry (PAT), combined with an increase in pulse rate, will accurately reflect arousals from sleep and can provide an autonomic arousal index (AAI). According to the authors, a previously studied group of 40 sleep apnea patients simultaneously recorded by both PSG and PAT systems generated an automated algorithm using the PAT signal (and pulse rate derived from it) was developed for detection of arousals from sleep. This was further validated in this separate group of 96 subjects which included 85 patients referred with suspected obstructive sleep apnea and 11 healthy volunteers. All subjects underwent a whole night PSG with simultaneous PAT recording. The PSG recordings were manually (blindly) analyzed for arousals based on American Academy of Sleep Medicine (AASM) criteria, while PAT was scored automatically. There was a significant correlation between PSG and PAT arousals (R=0.82, p<0.0001) with good agreement across a wide range of values, and with a ROC curve having an area under the curve (AUC) of 0.88. The authors conclude that automated analysis of the peripheral arterial tonometry signal can detect EEG arousals from sleep in a relatively quick and reproducible fashion.

Bar et al. (2003) aimed at evaluating the efficacy, reliability, and reproducibility of the Watch-PAT100 device for the diagnosis of OSAS as compared to in-laboratory, standard PSG-based manual scoring. One hundred two subjects (69 patients with OSAS and 33 normal non-consecutively selected volunteers) underwent in-laboratory full PSG simultaneously with Watch-PAT100 recording. Fourteen subjects also underwent two additional unattended home sleep studies with the Watch-PAT100 alone. The PSG recordings were blindly scored for apnea/hypopnea according to the American Academy of Sleep Medicine criteria (1999) and the RDI [PSG-RDI] was calculated. The Watch-PAT100 data were analyzed automatically for the PAT RDI (PRDI) by a proprietary algorithm that was the authors reported was previously developed on an independent group of subjects. Across a wide range of RDI levels, the PRDI was highly correlated with the PSG-RDI (r = 0.88, p < 0.0001), with an area under the receiver operating characteristic curve of 0.82 and 0.87 for thresholds of 10 events per hour and 20 events per hour, respectively. The PRDI scores were also highly reproducible, showing high correlation between home and in-laboratory sleep studies (r =0.89, p < 0.001). The authors concluded that the Watch-PAT100 may offer an accurate, robust, and reliable ambulatory method for the detection of OSAS with minimal patient discomfort.

Ayas et al. (2003) aimed at assessing the accuracy of a wrist-worn device (Watch-PAT100) to diagnose obstructive sleep apnea (OSA). Thirty adult subjects (mean age was 47.0 ± 14.8 years, mean body mass index 31.0 ± 7.6 kg/m2) were recruited through advertisements and from a patient base of those with suspected OSA to participate in this study. The study included patients suspected of having sleep apnea and subjects without suspected sleep apnea. The subjects had simultaneous in-laboratory PSG and wore the Watch-PAT 100 during a full-night recording. PSG sleep and respiratory events were scored according to standard criteria. The mean PSG AHI was 23 ± 23.9 events per hour and the mean PAT AHI 23 ± 15.9 events per hour. There was a significant correlation between the two (r = 0:87, p <0:001). To assess sensitivity and specificity of Watch-PAT100, receiver operator characteristic curves were constructed using a variety of AHI threshold values (10, 15, 20, and 30 events per hour). Optimal combinations of sensitivity and specificity for the various thresholds were 82.6/71.4, 93.3/73.3, 90.9/84.2, and 83.3/91.7, respectively. The authors concluded that the Watch-PAT100 is a device that can detect OSA with reasonable accuracy and that it may be a useful method to diagnose OSA.

Pillar et al. (2003) stated that they had recently shown that automated analysis of in-lab recorded peripheral arterial tone (PAT) signal and the pulse rate derived from it can accurately assess arousals from sleep as defined by the AASM. In the current study they aimed at extending these findings to the Watch-PAT100. They recruited 68 subjects who underwent a whole night PSG with simultaneous recording of PAT signal by the ambulatory Watch-PAT100 device. The PSG recordings were blindly scored via manual analyzing for arousals based on AASM criteria, while PAT was scored automatically based on the algorithm developed previously. The authors determined that was a significant correlation between AASM arousals derived from the PSG and PAT autonomic arousals derived from the Watch-PAT100 (R=0.87, P<0.001), with consistency across a wide range of values of AHI. The sensitivity and specificity of PAT in detecting patients with at least 20 arousals per hour of sleep were 0.80 and 0.79, respectively, with a receiver operating characteristic curve having an area under the curve of 0.87. They concluded that that automatic analysis of peripheral arterial tonometry signal derived from the ambulatory device Watch-PAT100 can accurately identify arousals from sleep in a simple and time saving fashion.

Berry et al. (2008) aimed to compare portable monitoring (PM) for diagnosis of OSA using the Watch PAT100 and unattended autotitrating positive airway pressure (APAP) for selecting an effective continuous positive airway pressure (CPAP), with polysomnography (PSG) for diagnosis and treatment of obstructive sleep apnea (OSA). The study was structured as a randomized parallel group comparison in a VA Medical Center. One hundred six patients with daytime sleepiness and a high likelihood of having OSA were recruited. The AHI in the PM-APAP group was 29.2 ± 2.3/h and in the PSG group was 36.8 ± 4.8/h (P = NS). Patients with an AHI ≥ 5 were offered CPAP treatment. Those accepting treatment (PM-APAP#= 45, PSG# = 43) were begun on CPAP using identical devices at similar mean pressures (11.2 ± 0.4 versus 10.9 ± 0.5 cm H2O).

At a clinic visit 6 weeks after starting CPAP, 40 patients in the PM-APAP group (78.4% of those with OSA and 88.8% started on CPAP) and 39 in the PSG arm (81.2% of those with OSA and 90.6% of those started on CPAP) were using CPAP treatment (P = NS). The mean nightly adherence (PM-APAP= 5.20 ± 0.28 hours/night versus PSG= 5.25 ± 0.38 hours/night), decrease in Epworth Sleepiness Scale score (–6.50 ± 0.71 versus –6.97 ± 0.73 in the PM group as compared to the PSG group respectively), improvement in the global Functional Outcome of Sleep Questionnaire score (3.10 ± 0.05 versus 3.31 ± 0.52 in the PM group as compared to the PSG group respectively), and CPAP satisfaction did not differ between the groups. The authors concluded that PM with APAP titration resulted in CPAP adherence and clinical outcomes similar to a diagnosis and treatment plan using PSG.

Other Diagnostic Strategies

Rice et al. (2006) piloted a study to evaluate unattended cardiopulmonary (CP) sleep studies as a diagnostic and treatment tool for patients with OSA. After all 106 subjects were initially evaluated by a pulmonary physician to identify those with a high risk of OSA, an ESS was administered. Those who were felt to have a high suspicion of OSA were offered either a PSG (which could take up to 6 months to schedule), or an unattended CP sleep study. Patients electing to use the unattended CP sleep study were lodged as outpatients overnight in the medical center. The diagnostic portable system used was the Embletta PDS, which included an oral thermometer, a nasal flow sensor, a snore microphone, a pulse oximeter, and strain gauges for thoracic and abdominal expansion. AHI was the outcome of interest. Patients with a positive CP test (an AHI of 5 events per hour or greater) were sent home with a REMstar auto CPAP system and a mask that was custom-fitted by a trained respiratory therapist.

After using auto CPAP nightly for a week (REMstar auto CPAP system adjusted to the patient's pressure needs by analyzing the shape curve of his/her airflow signal and peak flow), patients were then issued a home CPAP machine with settings based on the pressure that was found to be effective for at least 90% of the trial patients. ESS scores were measured at baseline and after 6 months of home CPAP use. Patients who had been prescribed home CPAP were assessed for global sleepiness at 12 months. CP studies were performed on 106 patients, all participants were males (mean age 59.9±10.1), mean BMI of 33.5 and mean ESS score (reference) of 13.1 ± 5.2. Of the 106 original patients, auto CPAP was initiated on 92 subjects. Based on the results of the one week auto CPAP, home CPAP was initiated on 84 patients. According to the authors, "among our patients, improvement in OSA symptoms and long-term adherence to prescribed CPAP was similar to published reports of patients who had undergone conventional PSG testing." At 6 months follow-up, 98% of CPAP patients were available; ESS scores at baseline and follow up were 14±4.6 and 10±5.6 (p=0.001), and adherence to CPAP usage was 84%.

Limitations of the study included the lack of confirmatory PSG to determine rate of false positives (but the mean AHI from this study was similar to that reported in published series of patients who had PSGs; and the absolute magnitude of ESS score improvement in this study was similar to that reported for patients who were prescribed CPAP after a PSG). Other limitations are the inability to calculate the diagnostic accuracy of a negative CP study for OSA; the fact that all subjects were male; and that adherence to prescribed CPAP was not based on objective data but rather on self-reporting.

4. MEDCAC

CMS convened the MEDCAC on September 12, 2007 to consider questions pertinent to the CPAP NCD reconsideration. We believe that many of those questions are relevant to this NCD consideration on sleep testing for OSA so we are reiterating them here. The questions are described below in reference to PSG and HST technologies. Additional information about the meeting can be found at:
https://www.cms.hhs.gov/mcd/viewmcac.asp?where=index&mid=40.

The MEDCAC was asked to consider the questions below for a variety of technologies including PSG, HST, clinical examination alone, and trial by CPAP without antecedent sleep testing.

1. How confident are you that there is sufficient evidence to determine if each of the following strategies can, in routine use, produce an accurate diagnosis of OSA for the prescription of CPAP?

2. For each OSA diagnostic strategy for which there is enough evidence in question 1, how confident are you about its sensitivity (ability to minimize false negatives) and specificity (ability to minimize false positives)?

3. How should each of the following factors be weighed as criteria for the prescription of CPAP for the diagnosis of OSA?

4. CPAP is currently a standard treatment for OSA. Defining successful treatment as combined subjective improvement of OSA clinical signs/symptoms and continued patient use of CPAP for 2 or more months, how confident are you that there is sufficient evidence to determine the ability of each of the following diagnostic strategies to accurately predict successful treatment of OSA with CPAP?

5. How confident are you that each of the following diagnostic strategies will accurately predict successful treatment of OSA with CPAP?

6. How confident are you that no clinically meaningful harm to patients will be caused by a Trial by CPAP strategy as an alternative to strategies that require a positive prior PSG or home sleep test before CPAP?

7. How confident are you that your conclusions can be generalized to the Medicare population and to providers in community practice?

The MEDCAC expressed moderate to high confidence that there was sufficient evidence to determine whether clinical evaluation combined with PSG can, in routine use, produce an accurate diagnosis of OSA for the prescription of CPAP treatment. Considering the evidence on the ability of various diagnostic strategies to predict successful use of CPAP, the MEDCAC expressed moderately high confidence in clinical evaluation combined with PSG; and moderate confidence in clinical evaluation combined with home sleep testing.

5. Evidenced based Guidelines

We identified the following evidence based guidelines that address the diagnosis of OSA.

American Academy of Sleep Medicine
Clinical guidelines for the use of unattended portable monitors in the diagnosis of OSA in adult patients. Portable Monitoring Task Force of the American Academy of Sleep Medicine. Collop NA, Anderson WM, Boehlecke B, Claman D, Goldberg R, Gottlieb DJ, Hudgel D, Sateia M, Schwab R; Portable Monitoring Task Force of the American Academy of Sleep Medicine. J Clin Sleep Med. 2007 Dec 15;3(7):737-47.

Based on a review of literature and consensus, the Portable Monitoring Task Force of the American Academy of Sleep Medicine (AASM) makes the following recommendations: unattended portable monitoring (PM) for the diagnosis of OSA (OSA) should be performed only in conjunction with a comprehensive sleep evaluation. Clinical sleep evaluations using PM must be supervised by a practitioner with board certification in sleep medicine or an individual who fulfills the eligibility criteria for the sleep medicine certification examination. PM may be used as an alternative to polysomnography (PSG) for the diagnosis of OSA in patients with a high pretest probability of moderate to severe OSA. PM is not appropriate for the diagnosis of OSA in patients with significant comorbid medical conditions that may degrade the accuracy of PM. PM is not appropriate for the diagnostic evaluation of patients suspected of having comorbid sleep disorders. PM is not appropriate for general screening of asymptomatic populations. PM may be indicated for the diagnosis of OSA in patients for whom in-laboratory PSG is not possible by virtue of immobility, safety, or critical illness. PM may also be indicated to monitor the response to non-CPAP treatments for sleep apnea. At a minimum, PM must record airflow, respiratory effort, and blood oxygenation. The airflow, effort, and oximetric biosensors conventionally used for in-laboratory PSG should be used in PM. The Task Force recommends that PM testing be performed under the auspices of an AASM-accredited comprehensive sleep medicine program with written policies and procedures. An experienced sleep technologist/technician must apply the sensors or directly educate patients in sensor application. The PM device must allow for display of raw data with the capability of manual scoring or editing of automated scoring by a qualified sleep technician/technologist. A board certified sleep specialist, or an individual who fulfills the eligibility criteria for the sleep medicine certification examination, must review the raw data from PM using scoring criteria consistent with current published AASM standards. Under the conditions specified above, PM may be used for unattended studies in the patient's home. A follow-up visit to review test results should be performed for all patients undergoing PM. Negative or technically inadequate PM tests in patients with a high pretest probability of moderate to severe OSA should prompt in-laboratory polysomnography.

Institute for Clinical Systems Improvement (ICSI)
Diagnosis and treatment of obstructive sleep apnea. Bloomington (MN): 2007 Mar. 55 p. [115 references]
http://www.guidelines.gov/summary/summary.aspx?doc_id=10809&nbr=005634&string=CPAP

Sleep Study
Key Points:

- Selection of appropriate diagnostic tests must take into account the estimated pretest probability of the patient having OSAHS, availability of credible diagnostic tests and local expertise in interpreting these tests.
- Polysomnography is the accepted standard test for the diagnosis of OSAHS.
- The benefit of using attended polysomnography for diagnosis is the ability to establish a diagnosis and ascertain an effective continuous PAP (CPAP) treatment pressure.

- Unattended portable recording (multichannel) is a second-best option for patients who have a high pretest probability of OSAHS and who do not have atypical or complicating symptoms.

Scottish Intercollegiate Guidelines Network (SIGN).
Management of obstructive sleep apnoea/hypopnoea syndrome in adults. A national clinical guideline. Edinburgh (Scotland): 2003 Jun. 35 p. (SIGN publication; no. 73). [158 references]. http://www.guidelines.gov/summary/summary.aspx?doc_id=3878&nbr=003087&string=CPAP
The SIGN guideline has the following recommendations including a letter grade for the grade of the recommendation:

A. *At least one meta-analysis, systematic review of randomised controlled trials (RCTs), or RCT rated as 1++ and directly applicable to the target population; or A body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results*
B. *A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+*
C. *A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++*
D. *Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+*

Diagnosis
C - All patients who have suspected sleep apnoea and their partners should complete an Epworth questionnaire to subjectively assess the degree of pretreatment sleepiness.

Diagnostic Tools
B - Limited sleep studies to assess respiratory events are an adequate first-line method of diagnostic assessment for obstructive sleep apnoea/hypopnoea syndrome (OSAHS).

American Society of Anesthesiologists
Practice guidelines for the perioperative management of patients with obstructive sleep apnea: a report by the Task Force on Perioperative Management of Patients with Obstructive Sleep Apnea. Anesthesiology 2006 May ;104(5):1081-93. [3 references].

Preoperative Evaluation

Anesthesiologists should work with surgeons to develop a protocol whereby patients in whom the possibility of obstructive sleep apnea (OSA) is suspected on clinical grounds are evaluated long enough before the day of surgery to allow preparation of a perioperative management plan. This evaluation may be initiated in a preanesthesia clinic (if available) or by direct consultation from the operating surgeon to the anesthesiologist. A preoperative evaluation should include a comprehensive review of previous medical records (if available), an interview with the patient and/or family, and conducting a physical examination. Medical records review should include (but not be limited to) checking for a history of airway difficulty with previous anesthetics, hypertension or other cardiovascular problems, and other congenital or acquired medical conditions. Review of sleep studies is encouraged. The patient and family interview should include focused questions related to snoring, apneic episodes, frequent arousals during sleep (vocalization, shifting position, extremity movements), morning headaches, and daytime somnolence. A physical examination should include an evaluation of the airway, nasopharyngeal characteristics, neck circumference, tonsil size, and tongue volume. If any of these characteristics suggest that the patient has OSA, the anesthesiologist and surgeon should jointly decide whether to (1) manage the patient perioperatively based on clinical criteria alone or (2) obtain sleep studies, conduct a more extensive airway examination, and initiate indicated OSA treatment in advance of surgery. If this evaluation does not occur until the day of surgery, the surgeon and anesthesiologist together may elect for presumptive management based on clinical criteria or a last-minute delay of surgery. For safety, clinical criteria (see table 1 of the original Guideline document) should be designed to have a high degree of sensitivity (despite the resulting low specificity), meaning that some patients may be treated more aggressively than would be necessary if a sleep study were available.

The severity of the patient's OSA, the invasiveness of the diagnostic or therapeutic procedure, and the requirement for postoperative analgesics should be taken into account in determining whether a patient is at increased perioperative risk from OSA (see table 2 of the original Guideline document). The patient and his or her family as well as the surgeon should be informed of the potential implications of OSA on the patient's perioperative course

University of Texas, School of Nursing, Family Nurse Practitioner Program.
Screening for obstructive sleep apnea in the primary care setting.
2006 May. 13 p. [24 references]
http://www.guidelines.gov/summary/summary.aspx?doc_id=9436&nbr=005057&string=
Diagnostic Procedures

1. Laboratory studies
    o   Sleep questionnaire (e.g., Epworth Sleepiness Scale), screen for sleep abnormalities (Elliott, 2001) (Strength of Recommendation: A; Quality of Evidence: Good)
2. Diagnostic tests
    o   NPSG Sleep Study: Nocturnal polysomnographic diagnostic testing (Netzer et al., 2003; Schroder, 2005; Elliot, 2001; Mansfield & Naughton, 2005; Hamilton, Solin, & Naughton, 2004; Rodsutti et al., 2004) (Strength of Recommendation: A; Quality of Evidence: Good)

6. Professional Society Position Statements

We expect to receive professional society position statements on this proposed decision.

7. Expert Opinion

We expect to receive expert opinion on this proposed decision.

8. Public Comments

As we are posting this proposed decision concurrently with opening the decision, we do not have an initial 30 day comment period upon opening an NCD. Posting this proposed decision begins the statutorily required comment period on the proposed decision and we will respond to those comments in the final decision.

VIII. CMS Analysis

National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally by Medicare (§1869(f)(1)(B) of the Act). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions the expenses incurred for items or services must be "reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member." See §1862(a)(1)(A)of the Social Security Act.

The Medicare regulations at 42 CFR 410.32(a) state in part, that "…diagnostic tests must be ordered by the physician who is treating the beneficiary, that is, the physician who furnishes a consultation or treats a beneficiary for a specific medical problem and who uses the results in the management of the beneficiary's specific medical problem."

As a diagnostic test, the sleep test would not be expected to directly change health outcomes, i.e. there is no evidence that administration of a sleep test is, in and of itself, therapeutic. Rather, a diagnostic test affects health outcomes through changes in disease management brought about by physician actions taken in response to test results. Such actions may include decisions to treat or withhold treatment, to choose one treatment modality over another, or to choose a different dose or duration of the same treatment. To some extent the usefulness of a test result is constrained by the available management alternatives.

Based on the legal framework set forth above, this section presents the agency's evaluation of the evidence considered and conclusions reached for the assessment questions posed above.

We considered the evidence in the hierarchical framework of Fryback and Thornbury (1991) where Level 2 addresses diagnostic accuracy, sensitivity, and specificity of the test; Level 3 focuses on whether the information produces change in the physician's diagnostic thinking; Level 4 concerns the effect on the patient management plan and Level 5 measures the effect of the diagnostic information on patient outcomes. Many studies have focused on test characteristics but others have considered health outcomes, such as symptom improvement in patients who receive CPAP treatment based on sleep test results. We believe that evidence of improved health outcomes is more persuasive than evidence of test characteristics.

In evaluating diagnostic tests, Mol and colleagues (2003) reported: "Whether or not patients are better off from undergoing a diagnostic test will depend on how test information is used to guide subsequent decisions on starting, stopping, or modifying treatment. Consequently, the practical value of a diagnostic test can only be assessed by taking into account subsequent health outcomes." When a proven, well established association or pathway is available, intermediate health outcomes may also be considered. For example, if a particular diagnostic test result can be shown to change patient management and other evidence has demonstrated that those patient management changes improve health outcomes, then those separate sources of evidence may be sufficient to demonstrate positive health outcomes from the diagnostic test.

A number of issues have emerged during our review of the evidence. Some physicians express concern about the lack of timely access to PSGs while others argue that access is not problematic; stakeholders debate the comparable accuracy of home apnea monitoring and PSGs; and recent research suggests that neither PSG nor HST may be needed for the diagnosis of CPAP-responsive OSA in selected patient populations. Paradoxically, some patients' OSA symptoms may be so severe that they cannot sleep for a sufficient continuous duration to complete a PSG. Adding to the complexity of our review, the stakeholder community itself has been clearly polarized into opposing PSG and HST camps.

The relevance of OSA diagnosis is founded on the long term morbidity and mortality that have been observed in patients who display a particular constellation of symptoms, signs and test results. Absent that morbidity and mortality, a self-limited apneic episode in and of itself appears of little consequence. Hence the challenge is to select for treatment only those patients who will benefit from OSA therapy, against a background of persons who may for various reasons have a normal or abnormal test on a given night.

It is important to state at the outset that sleep testing, whether via PSG or HST, is used to confirm or refute a clinical suspicion of OSA. In other words, we have no evidence that physicians refer "normal" patients, i.e. patients who manifest no symptoms or signs of a sleep disorder, for sleep testing. Sleep testing does not occur in a vacuum, divorced from the overall clinical evaluation. We also note that this NCA speaks to sleep testing for the diagnosis of OSA; we are not establishing coverage criteria for the diagnosis of other sleep disorders, such as nocturnal seizures or restless legs syndrome. Hence the accurate identification of Medicare beneficiaries who have OSA is at the heart of this review and analysis.

PSG is utilized as a reference standard in many clinical trials; however, we do not believe it is a true gold standard. In a circular argument, the test result has been incorporated into the diagnosis of the disease itself. In the absence of a diagnostic gold standard, this is an understandable though not ideal concession to practicality. The accuracy and precision of PSG may be compromised by many factors such as inter-reader variability, the use of different test instruments, night to night variability in a given patient, and patient ability to sleep in a non-home setting. Even if all these variables are controlled, the PSG test itself has not been proven to identify all true cases of OSA, i.e. those persons who will develop OSA-associated morbidity and mortality if untreated.

Therefore, when PSG is performed and read with a threshold of AHI $\geq$ 15 events per hour for OSA, the sensitivity for detecting a true case of OSA is not known. Neither is its specificity for detecting those who do not have OSA truly known. An AHI suggestive of OSAHS does not conclusively identify those patients who will benefit from treatment. Since the true sensitivity and specificity of PSG are uncertain and the reported agreement between HST and PSG is not complete, we are concerned that some true cases of OSA are not detected by either test. Nonetheless, it is the current state of the art and we believe that the evidence is sufficient to conclude that, despite the lack of agreement on a true gold standard for diagnosis, these sleep test technologies do provide useful clinical information.

Question 1: Is the evidence adequate to determine that attended facility based polysomnography accurately identifies patients with OSA who will benefit from treatment?

Attended facility based PSG for this use is well supported by evidence based guidelines and, though imperfect as noted above, is the generally accepted reference standard for the diagnosis of OSA. The external TA is consistent with our own review of the evidence and supports this conclusion.

The reference standard for the diagnosis of OSAHS is facility-based polysomnography (PSG), a comprehensive sleep study that records and evaluates a variety of cardiorespiratory and neurophysiologic signals during sleep time. It quantifies the severity of disturbances with the Apnea-Hypopnea Index (AHI).

Based on our prior and current review of the evidence we believe that PSG accurately identifies patients with OSA who will benefit from treatment and therefore improve health outcomes by identifying patients with OSA who are likely to benefit from CPAP therapy. The external TA is consistent with our own review of the evidence and supports this conclusion.

The mainstay of treatment is considered to be continuous positive airway pressure (CPAP). Other treatments for the condition exist and are reserved for specific cases (e.g., surgical interventions and oral-dental appliances to improve the stereometry of the upper airway).

CPAP treatment of OSAHS has been associated with beneficial health outcomes. Observational evidence from prospective comparative studies associates CPAP treatment of OSAHS with fewer cardiovascular events. Furthermore, patients with OSAHS have an increased risk for car accidents. CPAP has been associated with a reduction in the risk for motor vehicle accidents among people with OSAHS.

However, apart from the aforementioned considerations there is no extensive randomized evidence on outcomes such as deaths, strokes and cardiovascular events. There is randomized evidence that CPAP versus no treatment or sham CPAP treatment of OSAHS is associated with improvements in the Epworth Sleepiness Scale (a subjective symptom score), objective wakefulness tests and selected components of the SF-36 questionnaire (e.g., the vitality component, which is more relevant to OSAHS patients compared to other SF-36 components). Randomized studies suggest that CPAP may also be inversely associated with intermediate clinical outcomes (e.g., hypertension).

Typically, the diagnosis of OSAHS is made after a positive comprehensive sleep study with multichannel polysomnography (PSG) in specialized sleep laboratories. For patients who meet the diagnostic criteria, a second session is needed for the titration

Therefore, we propose that PSG is reasonable and necessary for the diagnosis of OSA.

Question 2: For which unattended out of facility sleep test technologies is the evidence adequate to determine that sleep testing accurately identifies patients with OSA who will benefit from treatment?

We note evidence from our internal assessment and from the AHRQ TAs that HST devices may, with high positive likelihood ratios (> 10) and low negative likelihood ratios (< 0.1), identify patients who have AHIs suggestive of OSAHS. Although there is published data comparing HST with PSG, in the absence of a true gold standard it is challenging to categorize the discrepancies as errors.

The body of evidence pertinent to the use of HST devices for the diagnosis of OSA is significantly more robust than it was a few years ago. This is supported by the more favorable September 2007 MEDCAC scores for HST compared to the September 2004 MCAC scores. Thus, we find that the evidence is sufficient to conclude that, in appropriately selected patients, some home sleep testing monitors will identify a significant proportion of patients with OSA who will respond clinically to CPAP and will exclude a significant proportion of those who will not.

Specifically, we believe that Type II and Type III sleep testing devices, based on our prior and current review of the evidence, identify beneficiaries with OSA who will benefit from treatment and thus improve health outcomes by identifying patients who are likely to respond to CPAP therapy as we discussed above. Therefore, we propose that Type II and Type III sleep testing devices are reasonable and necessary for the diagnosis of OSA.

The TA analyzed Type IV monitors with three or more channels separately from those with only one or two channels. The quality of the evidence on the former is described as higher than on the latter. We also note that the TA did not include all Type IV monitors.

"…However, especially for type IV devices, we excluded the few studies that did not measure directly at least one respiratory signal or the O2 saturation. Thus, studies using only static charge-sensitive mattresses, only Holter recordings for heart rate, or studies that used only analysis of snoring sounds were excluded. Similarly, we excluded studies that that used pulse oximetry but analyzed only the variability of the heart rate (i.e., used oximetry in lieu of ECG to detect pulse rate) and did not evaluate O2 saturation patterns. In general, monitors that did not record a respiratory signal or SaO2 during sleep rely on "indirect" assessment of respiratory disturbances in people suspected for OSAHS, and most often were described in older studies. The frequency of respiratory disturbances is a key issue in the diagnosis of OSAHS, and is assessed by the vast majority of modern monitors."

Based on this, as well as our prior and current review of the evidence, we believe Type IV sleep testing devices that measure three or more channels, one of which is airflow, identify beneficiaries with OSA who will benefit from treatment and thus improve health outcomes. Therefore, we propose that Type IV sleep testing devices that measure three or more channels, one of which is airflow, are reasonable and necessary for the diagnosis of OSA.

We have separately reviewed evidence concerning sleep testing devices measuring three or more channels that include actigraphy, oximetry, and peripheral arterial tone. We believe that the evidence is sufficient for us to confidently conclude that such devices identify beneficiaries with OSA who will benefit from treatment and thus improve health outcomes. Therefore, we propose that sleep testing devices measuring three or more channels that include actigraphy, oximetry, and peripheral arterial tone are reasonable and necessary for the diagnosis of OSA.

X. Proposed Conclusion

CMS proposes that the evidence is sufficient to determine that the results of the sleep tests identified below can be used by a beneficiary's treating physician to diagnose OSA and prescribe CPAP therapy, that the use of such sleep testing technologies demonstrates improved health outcomes in Medicare beneficiaries who have OSA and receive the appropriate treatment, and that these tests are thus reasonable and necessary under section 1862(a)(1)(A) of the Social Security Act.

Therefore, we propose that:

1.

    Type I Polysomnography (PSG) is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed attended in a sleep lab facility.

2.

>A Type II or a Type III sleep testing device is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have clinical signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

3.

>A Type IV sleep testing device measuring three or more channels, one of which is airflow, is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

4.

>A sleep testing device measuring three or more channels that include actigraphy, oximetry, and peripheral arterial tone is covered when used to aid the diagnosis of obstructive sleep apnea (OSA) in beneficiaries who have signs and symptoms indicative of OSA if performed unattended in or out of a sleep lab facility or attended in a sleep lab facility.

We are soliciting public comments on this proposed decisions pursuant to §1862(l) of the Social Security Act.

APPENDIX A

General Methodological Principles of Study Design
(Section VI of the Decision Memorandum)

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service is reasonable and necessary. The overall objective for the critical appraisal of the evidence is to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients.

We divide the assessment of clinical evidence into three stages: 1) the quality of the individual studies; 2) the generalizability of findings from individual studies to the Medicare population; and 3) overarching conclusions that can be drawn from the body of the evidence on the direction and magnitude of the intervention's potential risks and benefits.

The methodological principles described below represent a broad discussion of the issues we consider when reviewing clinical evidence. However, it should be noted that each coverage determination has its unique methodological aspects.

Assessing Individual Studies

Methodologists have developed criteria to determine weaknesses and strengths of clinical research. Strength of evidence generally refers to: 1) the scientific validity underlying study findings regarding causal relationships between health care interventions and health outcomes; and 2) the reduction of bias. In general, some of the methodological attributes associated with stronger evidence include those listed below:

- Use of randomization (allocation of patients to either intervention or control group) in order to minimize bias.
- Use of contemporaneous control groups (rather than historical controls) in order to ensure comparability between the intervention and control groups.
- Prospective (rather than retrospective) studies to ensure a more thorough and systematical assessment of factors related to outcomes.
- Larger sample sizes in studies to demonstrate both statistically significant as well as clinically significant outcomes that can be extrapolated to the Medicare population. Sample size should be large enough to make chance an unlikely explanation for what was found.

- Masking (blinding) to ensure patients and investigators do not know to which group patients were assigned (intervention or control). This is important especially in subjective outcomes, such as pain or quality of life, where enthusiasm and psychological factors may lead to an improved perceived outcome by either the patient or assessor.

Regardless of whether the design of a study is a randomized controlled trial, a non-randomized controlled trial, a cohort study or a case-control study, the primary criterion for methodological strength or quality is the extent to which differences between intervention and control groups can be attributed to the intervention studied. This is known as internal validity. Various types of bias can undermine internal validity. These include:

- Different characteristics between patients participating and those theoretically eligible for study but not participating (selection bias).
- Co-interventions or provision of care apart from the intervention under evaluation (performance bias).
- Differential assessment of outcome (detection bias).
- Occurrence and reporting of patients who do not complete the study (attrition bias).

In principle, rankings of research design have been based on the ability of each study design category to minimize these biases. A randomized controlled trial minimizes systematic bias (in theory) by selecting a sample of participants from a particular population and allocating them randomly to the intervention and control groups. Thus, in general, randomized controlled studies have been typically assigned the greatest strength, followed by non-randomized clinical trials and controlled observational studies. The design, conduct and analysis of trials are important factors as well. For example, a well designed and conducted observational study with a large sample size may provide stronger evidence than a poorly designed and conducted randomized controlled trial with a small sample size. The following is a representative list of study designs (some of which have alternative names) ranked from most to least methodologically rigorous in their potential ability to minimize systematic bias:

Randomized controlled trials
Non-randomized controlled trials
Prospective cohort studies
Retrospective case control studies
Cross-sectional studies
Surveillance studies (e.g., using registries or surveys)
Consecutive case series
Single case reports

When there are merely associations but not causal relationships between a study's variables and outcomes, it is important not to draw causal inferences. Confounding refers to independent variables that systematically vary with the causal variable. This distorts measurement of the outcome of interest because its effect size is mixed with the effects of other extraneous factors. For observational, and in some cases randomized controlled trials, the method in which confounding factors are handled (either through stratification or appropriate statistical modeling) are of particular concern. For example, in order to interpret and generalize conclusions to our population of Medicare patients, it may be necessary for studies to match or stratify their intervention and control groups by patient age or co-morbidities.

Methodological strength is, therefore, a multidimensional concept that relates to the design, implementation and analysis of a clinical study. In addition, thorough documentation of the conduct of the research, particularly study selection criteria, rate of attrition and process for data collection, is essential for CMS to adequately assess and consider the evidence.

Generalizability of Clinical Evidence to the Medicare Population

The applicability of the results of a study to other populations, settings, treatment regimens and outcomes assessed is known as external validity. Even well-designed and well-conducted trials may not supply the evidence needed if the results of a study are not applicable to the Medicare population. Evidence that provides accurate information about a population or setting not well represented in the Medicare program would be considered but would suffer from limited generalizability.

The extent to which the results of a trial are applicable to other circumstances is often a matter of judgment that depends on specific study characteristics, primarily the patient population studied (age, sex, severity of disease and presence of co-morbidities) and the care setting (primary to tertiary level of care, as well as the experience and specialization of the care provider). Additional relevant variables are treatment regimens (dosage, timing and route of administration), co-interventions or concomitant therapies, and type of outcome and length of follow-up.

The level of care and the experience of the providers in the study are other crucial elements in assessing a study's external validity. Trial participants in an academic medical center may receive more or different attention than is typically available in non-tertiary settings. For example, an investigator's lengthy and detailed explanations of the potential benefits of the intervention and/or the use of new equipment provided to the academic center by the study sponsor may raise doubts about the applicability of study findings to community practice.

Given the evidence available in the research literature, some degree of generalization about an intervention's potential benefits and harms is invariably required in making coverage determinations for the Medicare population. Conditions that assist us in making reasonable generalizations are biologic plausibility, similarities between the populations studied and Medicare patients (age, sex, ethnicity and clinical presentation) and similarities of the intervention studied to those that would be routinely available in community practice.

A study's selected outcomes are an important consideration in generalizing available clinical evidence to Medicare coverage determinations. One of the goals of our determination process is to assess health outcomes. These outcomes include resultant risks and benefits such as increased or decreased morbidity and mortality. In order to make this determination, it is often necessary to evaluate whether the strength of the evidence is adequate to draw conclusions about the direction and magnitude of each individual outcome relevant to the intervention under study. In addition, it is important that an intervention's benefits are clinically significant and durable, rather than marginal or short-lived. Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits.

If key health outcomes have not been studied or the direction of clinical effect is inconclusive, we may also evaluate the strength and adequacy of indirect evidence linking intermediate or surrogate outcomes to our outcomes of interest.

Assessing the Relative Magnitude of Risks and Benefits

Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits. Health outcomes are one of several considerations in determining whether an item or service is reasonable and necessary. CMS places greater emphasis on health outcomes actually experienced by patients, such as quality of life, functional status, duration of disability, morbidity and mortality, and less emphasis on outcomes that patients do not directly experience, such as intermediate outcomes, surrogate outcomes, and laboratory or radiographic responses. The direction, magnitude, and consistency of the risks and benefits across studies are also important considerations. Based on the analysis of the strength of the evidence, CMS assesses the relative magnitude of an intervention or technology's benefits and risk of harm to Medicare beneficiaries.

Back to Top

Bibliography

Alonso Alvarez Mde L, Terán Santos J, Cordero Guevara J, et al. Reliability of Home Respiratory Polygraphy for the Diagnosis of Sleep Apnea-Hypopnea Syndrome. Analysis of Cost. Arch Bronconeumol. 2008 Jan;44(1):22-8

Alvarez M et al. Reliability of Home Respiratory Polygraphy for the Diagnosis of Sleep Apnea-Hypopnea Syndrome. Analysis of Costs. Arch Bronconeumol. 2008;44(1):22-8

American Sleep Disorder Association. Practice parameters for the indications of polysomnography and related procedures. Polysomnography Task Force, American Sleep Disorder Association Standards of Practice Committee. Sleep 1997;20:406-422.

American Thoracic Society. Indications and standards for use of nasal continuous positive airway pressure (CPAP) in sleep apnea syndrome-Official statement adopted March 1994. Am J Resp Crit Care Med. 1994;150:1738-1745.

Ayappa I, Norman RG, Seelall V, Rapoport DM. Validation of a Self-Applied Unattended Monitor for Sleep Disordered Breathing. J Clin Sleep Med. 2008 Feb 15;4(1):26-37.

Ayappa I, Norman RG, Suryadevara M, Rapoport DM. Comparison of limited monitoring using a nasal cannula flow signal to full polysomnography in sleep-disordered breathing. Sleep 2004;27(6):1171-1179.

Ayas NT et al. Assessment of a wrist-worn device in the detection of obstructive sleep apnea. Sleep Medicine 4 (2003) 435–442

Bar A et al. Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies. Chest 2003;123;695-703

Berry RB, Hill G, Thompson T, McLaurin V. Portable Monitoring and Autotitration versus Polysomnography for the Diagnosis and Treatment of Sleep Apnea. SLEEP 2008;31(10):1423-1431.

Berry RB, Parish JM, Hartse KM. The use of auto-titrating continuous positive airway pressure for treatment of adult obstructive sleep apnea. Sleep 2002;25(2):148-173.

Caples SA, Gami AS, Somers VK. Obstructive Sleep Apnea. Annals of Internal Medicine 2005;142:187-197.

Chiner E, Signes-Costa J, Arriero J, Marco J, Fuentes I, Sergado A. Nocturnal oximetry for the diagnosis of the sleep apnoea hypopnoea syndrome: a method to reduce the number of polysomnographies. Thorax 1999;54:968-971.

Claman D, Murr A, Trotter K. Clinical validation of the Bedbug™ in detection of obstructive sleep apnea. Otolaryngology-Head and Neck Surgery 2001;125(3):227-230.

Crocker BD et al. Estimation of the probability of disturbed breathing during sleep before a sleep study. Am Rev Respir Dis. 1990 Jul;142(1):14-8.

Deegan PC et al. Predictive value of clinical features for the obstructive sleep apnoea syndrome. Eur Respir J. 1996 Jan;9(1):117-24

Dixon J et al. Predicting sleep apnea and excessive day sleepiness in the severely obese: indicators for polysomnography. Chest. 2003 Apr;123(4):1134-41.

Fitzpatrick M, Alloway C, Wakeford T, MacLean A, Munt P, Day A. Can patients with obstructive sleep apnea titrate their own continuous positive airway pressure? Am J Respir Crit Care Med 2003;167:716-722.

Flemons WW, Whitelaw WA, Brant R, Remmers JE. Likelihood ratios for a sleep apnea clinical predictive rule. Am Jnl of Respi Crit Care Med. 1994;150:1279-1285.

Garcia MA et al.Clinical Predictors of Sleep Apnea-Hypopnea Syndrome Susceptible to Treatment With Continuous Positive Airway Pressure. Arch Bronconeumol 2003;39(10):449-54

Golpe R, Jimenez A, Carpizo R, Cifrian JM, Utility of home oximetry as a screening test for patients with moderate to severe symptoms of obstructive sleep apnea. Sleep 1999 Nov 1;22(7):932-937.

Golpe R, Jimenez A, Carpizo R. Home sleep studies in the assessment of sleep apnea/hyponea syndrome. Chest 2002;122:1156-1161.

Gyulay S, Olson LG, Hensley MJ, King MT, Allen KM, Saunders NA. A comparison of clinical assessment and home oximetry in the diagnosis of obstructive sleep apnea. Am. Rev Respir Dis. 1993 Jan;147(1):50-53.

Haponik EF et al. Evaluation of sleep-disordered breathing. Is polysomnography necessary? Am J Med. 1984 Oct;77(4):671-7

Hoffstein et al. Predictive value of clinical features in diagnosing OSA .Sleep 1993;16:2:118-122

Hussain SF, Fleetham JA. Overnight home oximetry: can it identify patients with obstructive sleep apnea-hypopnea who have minimal daytime sleepiness. Respir Med. 2003 May;97(5):537-540.

Jenkinson C, Davies RJ, Mullins R, Stradling JR. Comparison of therapeutic and sub-therapeutic nasal continuous positive airway pressure for obstructive sleep apnea: a randomized prospective parallel trail. Lancet 1999;353:2100-2105

Julia-Serda G et al. Usefulness of cephalometry in sparing polysomnography. Sleep Breath (2006) 10:181–187

Kingshott RN, Vennelle M, Hoy CJ et al. Predictors of improvements in daytime function outcome with CPAP therapy. Am J Respir Crit Care Med 2000;161:866-877.

Kushida et al. A Predictive Morphometric Model for the Obstructive Sleep Apnea Syndrome. Ann Int Med 1997:127;8;1;581-587

Le Bon O, Hoffman G, Tecco J, et al. Mild to moderate sleep respiratory events: one night may not be enough. Chest 2000;118:353-359.

Lim PVH et al. The role of history, epworth sleepiness scale score and body mass index in identifying non-apnoeic snorers. Clin. Otolaryngol. 2000, 25, 244±248

Littner M. Polysomnography in the diagnosis of the obstructive sleep apnea-hypopnea syndrome: Where do we draw the line? Chest 2000;118:286-288.

Loube DI, Gay PC, Strohl KP, et al. Indications for positive airway pressure treatment of adults obstructive sleep apnea patients. Chest 1999;155:863-866.

Magalang, UJ, Dmochowski J, Veeramachaneji S, Draw A, Mador J, El-Solh A. Prediction of the apnea-hypopnea index from overnight pulse oximetry. Chest 2003;124:1694-1701.

Maislin G, Pack AI, Kribbs NB, Smith PL, Schwartz AR, Kline LR et al. A survey screen for prediction of sleep apnea. Sleep 1995;18:158-166.

Montserrat JM, Ferrer M, Hernandez L, Farre R, Vilagut G, Navajas D et al. Effectiveness of CPAP treatment in daytime function in sleep apnea syndrome: a randomized controlled study with an optimized placebo. American J Respir Care Med 2001;164;608-613.

Mulgrew A, Fox N, Ayas N, Ryan F. Diagnosis and initial management of obstructive sleep apnea without polysomnography. Annals of Internal Medicine 2007;146:157-166.

Nuber R, Vavrina J, Karrer W. Predictive value in nocturnal pulse oximetry in sleep apnea screening. Schweiz Med Wochenschr Suppl. 2000;116:120S-122S.
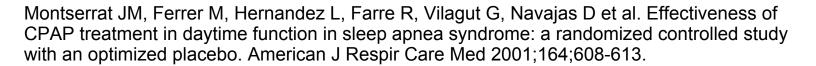
Pillar G, Bar A, Betito M, Schnall RP, Dvir I, Sheffy J, Lavie P. An automatic ambulatory device for detection of AASM defined arousals from sleep: the WP100. Sleep Med. 2003 May;4(3):207-12

Pillar G, Bar A, Shlitner A, Schnall R, Shefy J, Lavie P. Autonomic arousal index: an automated detection based on peripheral arterial tonometry. Sleep. 2002 Aug 1;25(5):543-9

Pillar G, Peled N, Katz N, Lavie P. Predictive value of specific risk factors, symptoms and signs, in diagnosing obstructive sleep apnoea and its severity. J Sleep Res. 1994 Dec;3(4):241-244

Pittman et al. Using a Wrist-Worn Device Based on Peripheral Arterial Tonometry to Diagnose Obstructive Sleep Apnea: In-Laboratory and Ambulatory Validation. SLEEP 2004;27(5):923-33

Rauscher H, Popp W, Zwick H. Model for investigating snorers with suspected sleep apnoea. Thorax 1993;48: 275–279.

Rice KL, Nelson K, Rubin JB, Arjes S. Unattended cardiopulmonary sleep studies to diagnose obstructive sleep apnea. Federal Practioner May 2006:17-31.

Rowley JA, Aboussouan LS, Badr MS. The use of clinical prediction formulas in the evaluation of obstructive sleep apnea. Sleep 2000;23:929-938.
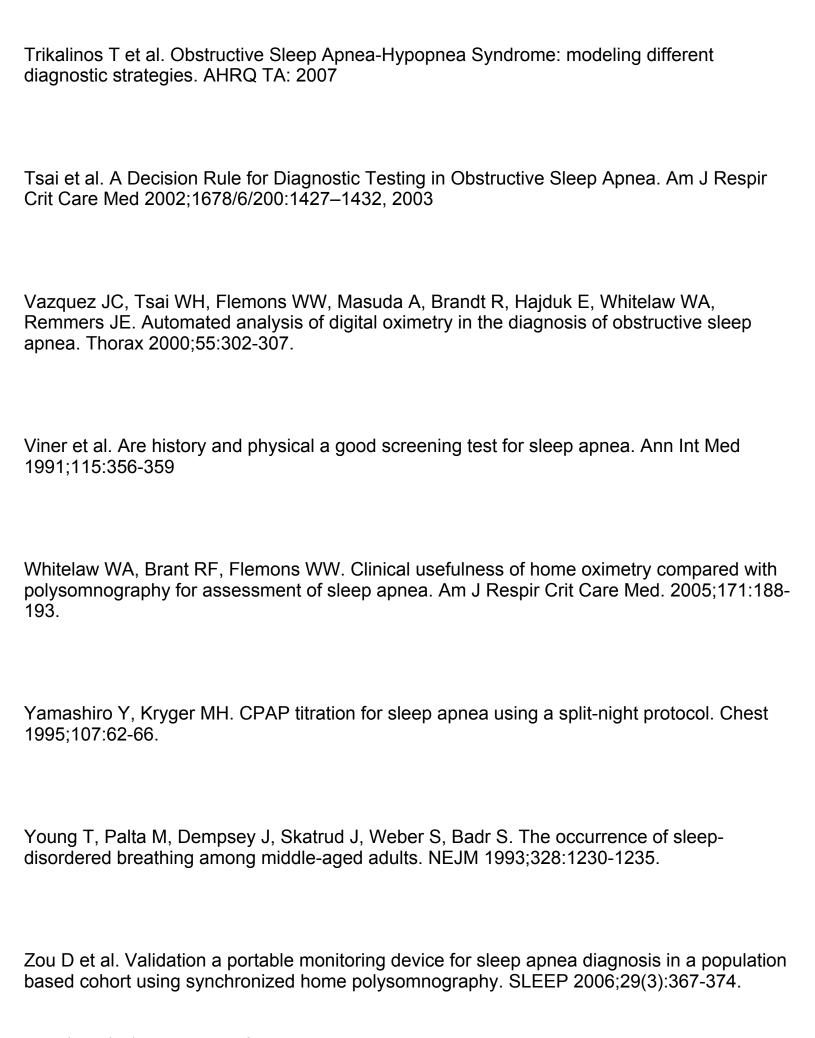
Sériés F, Kimoff RJ, Morrison D, Leblanc MH, Smilovitch M, Howlett J, et al. Prospective evaluation of nocturnal oximetry for detection of sleep-related breathing disturbances in patients with chronic heart failure. Chest 2005 May;127(5):1507-1514.

Sériés F, Marc I, Cormier Y, La Forge J. Utility of nocturnal home oximetry for case findings in patients with suspected sleep apnea syndrome. Ann Internal Med. 1993;119:449-453.

Sériés F, Marc I. Efficacy of automatic continuous positive pressure airway pressure therapy that uses an estimated required pressure in the treatment of obstructive sleep apnea syndrome. Ann Intern Med 1997;127:588-595.

Teran-Santos J, Jimenez-Gomez A, Cordero-Guevera J. The association between obstructive sleep apnea and the risk of traffic accidents. Cooperative Group Burgos Santander, NEJM 1993;340:847-851.

Trikalinos T et al. Home Diagnosis of Obstructive Sleep Apnea-Hypopnea Syndrome. AHRQ TA: 2007

Trikalinos T et al. Obstructive Sleep Apnea-Hypopnea Syndrome: modeling different diagnostic strategies. AHRQ TA: 2007

Tsai et al. A Decision Rule for Diagnostic Testing in Obstructive Sleep Apnea. Am J Respir Crit Care Med 2002;1678/6/200:1427–1432, 2003

Vazquez JC, Tsai WH, Flemons WW, Masuda A, Brandt R, Hajduk E, Whitelaw WA, Remmers JE. Automated analysis of digital oximetry in the diagnosis of obstructive sleep apnea. Thorax 2000;55:302-307.

Viner et al. Are history and physical a good screening test for sleep apnea. Ann Int Med 1991;115:356-359

Whitelaw WA, Brant RF, Flemons WW. Clinical usefulness of home oximetry compared with polysomnography for assessment of sleep apnea. Am J Respir Crit Care Med. 2005;171:188-193.

Yamashiro Y, Kryger MH. CPAP titration for sleep apnea using a split-night protocol. Chest 1995;107:62-66.

Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. NEJM 1993;328:1230-1235.

Zou D et al. Validation a portable monitoring device for sleep apnea diagnosis in a population based cohort using synchronized home polysomnography. SLEEP 2006;29(3):367-374.